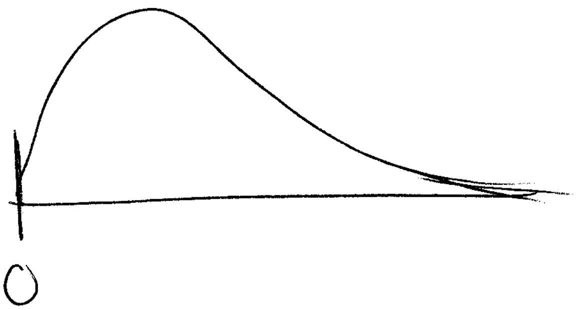
Key

Chapter 9 Review
Chi Squared Distributions

What does a X^2 distribution look like?

a. DRAW:



b. Describe the shape:

- rt. skewed
- center @ ?
- spread 0 → ∞

How many X^2 distributions are there?

infinitely many - based on sample size

Like the t-distribution, what do we ALWAYS have to include with the p-value for a X^2 test of significance?

df!

What is different about the hypotheses for a X^2 test of significance?

- written out
- full sentences
- no #'s
- always "two-sided" (no > or < alt.)

frequency! #'s

X^2 Goodness of Fit Test

What type of data is this test done on?

one-var. distribution (list of data) categories-

What are we trying to assess?

whether the observed distr. (from sample, expt) fits expected (from pop.)

Hypotheses: Ho: the observed freq. distribution of (context) fits the expected distribution

Ha: " " " " doesn't fit " " "

Test Statistic: $X^2 = \sum \frac{(obs-exp)^2}{exp} = \frac{(\quad)^2}{} + \frac{(\quad)^2}{} + \ldots =$ test stat

P-value: $P(\chi^2 >$ test stat $| df = \_\_\_) =$

df = $k-1$ (# categories − 1)

Conclusion: We reject/fail to reject….. Ho…

We have suff. evidence that… (re-copy Ho or Ha)

Assumptions: 1) SRS

2) large n so that all expected counts ≥ 5

On calculator:

$L_1 = obs$

$L_2 = exp$

$L_3 = (L_1 - L_2)^2 / L_2$

$sum(L_3) = \chi^2$

P-val: $\chi^2cdf(LB, UB, df)$

test stat E99

Example: #1

In a recent year, at the 6 P.M. time slot, television channels 3, 6, 10, and 29 captured the entire audience with 30%, 25%, 20%, and 25%, respectively. During the first week of the next season, 500 viewers are interviewed.

a. If viewer preferences have not changed, what number of persons is expected to watch each channel?

State
1) SRS
2) all expected counts ≥ 5

Check
1) assumed
2) ✓

Channel	exp	obs
3	150	139
6	125	138
10	100	112
29	125	111

b. If the actual observed values are 139, 138, 112, and 111, is there a significant difference?

Ho: the observed freq. distrib. of TV channels fits the expected distrib.

Ha: the observed freq. distrib. of TV channels doesn't fit the expected distrib.

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = \frac{(139-150)^2}{150} + \frac{(138-125)^2}{150} + \dots = 5.1667$$

$$P(\chi^2 > 5.1667 \mid df = 3) = 0.16$$

- We fail to reject Ho b/c p-value > $\alpha = 0.05$.
- We have sufficient evidence that.... (re-copy Ho).

X^2 Test for Association

What type of data is this test done on?

2 var. distribution (2 way table

What are we trying to assess?

whether obs fits expected distr.

How do we find the expected cell counts for a table of data?

$$\text{exp. counts} = \frac{\text{row total} \times \text{column total}}{n}$$

Hypotheses: **Ho:** there is no association btw. row var. and column var., or (they are indep)

Ha: there is an assoc. btw. .... or (they are dependent)

Test Statistic: $X^2 = \sum \frac{(obs - exp)^2}{exp} = \frac{(\quad)^2}{} + \frac{(\quad)^2}{} + \ldots = \text{test stat}$

P-value: **P(** $\chi^2 >$ test stat $|$ df $=$ $) =$

df = $(r-1)(c-1)$

Conclusion: **We reject/fail to reject.....**

We have suff. evidence that... re-copy Ho / Ha

Assumptions: **1)** 2 indep. SRS

2) large enough n so all exp. counts ≥ 5

On calculator:

- put data into matrix [A]
- Use χ^2 test
- expected are put into [B] by calculator

Example:

The following table is from the July 1993 publication of *Vital and Health Statistics* from the Centers for Disease Control and Prevention/National Center for Health Statistics. The individuals in the following table have only one of the three indicated irritations.

Irritation	18-29	30-44	45-64	65+	total
Eye	440	567	349	59	1415
Nose	924	1311	794	102	3131
Throat	253	311	157	19	740
total	1617	2189	1300	180	5286

State	Check
1) 2 indep. SRS	1) assumed
2) all exp. counts ≥ 5	2) ✓

a. How many rows are there in the table above?

3

b. How many cells are there?

12

c. How many columns are there?

4

d. What is the df?

$(3-1)(4-1) = 6$

e. Determine if the type of irritation is independent of the age group using a 0.05 level of significance.

Ho: irritation + age are independent

Ha: " " " dependent

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = \frac{(440-432.85)^2}{432.85} + \frac{(567-585.97)^2}{585.97} \ldots = 13.619$$

$P(\chi^2 \geq 13.619 \mid df = 6) = 0.034$

We reject Ho b/c p-value < $\alpha = 0.05$.

~~dependent~~

We have sufficient evidence that irritation + age are dependent.